PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A qualitative study of manufacturers' submissions to the UK NICE single technology appraisal process
AUTHORS	Kaltenthaler E, Dickson R, Boland A, Carroll C, Fitzgerald P, Diana Papaioannou and Ron Akehurst

VERSION 1 - REVIEW

REVIEWER	Professor Trevor Sheldon Deputy Vice-Chancellor University of York England
REVIEW RETURNED	21/11/2011

GENERAL COMMENTS	This paper reports a qualitative analysis of assessments of submissions from manufacturers to NICE as part of the Single Technology Appraisal process (STA) and letters of clarification to manufacturers in response to their submission. There has been concern about the STA process and whether NICE should base its reimbursement decisions exclusively on an assessment of information provided by manufacturers, instead of the original, more thorough, but time consuming process. This paper provides useful information about the quality of manufacturers' submissions as seen through reports and letters of the external, independent review groups. They find that there are significant variations in the quality of manufacturers' submissions and generally inadequate rigour and transparency in their contents, leading to concerns about bias. They present a set of 12 sensible recommendations for manufacturers to follow. The results are hardly surprising; companies often will not have the academic resources to carry out assessments to an appropriately high level and because of their obvious commercial interests they also have an incentive to present the information in the most favourable light. In addition, it will take time for companies to fully understand what is expected in a new process. The analysis was well conducted, the results reliable and the paper is clearly written. It is a useful, if not a particularly significant contribution to the literature. It would have been useful to know if the standard of submissions has been improving over time. Was there for example, any difference between those at the beginning of the process and those submitted further on? I was somewhat surprised that the researchers had to write these recommendations and send them to the manufacturers; this would

	appear to be a more appropriate role for NICE. Perhaps, I have misunderstood this. The research would have been strengthened if it has been accompanied by interviews with staff who have carried out the reviews and possibly some in NICE. It could usefully be followed up by research on the perceptions of the manufacturers to better understand how the quality of submissions could be improved. The authors could improve the paper by addressing the following points: 1) Shorten the objectives para in the Abstract (p3) so as to reduce the text on background – or add a background subheading 2) Add 'government' as a source of pressure for speedier assessment (page 4, line 20) 3) Change 'compactors' to 'comparators' (page 5, line 42) 4) Possibly better distinguish between bias and uncertainty (page 10, lines 26-36). Uncertainty is also, and often mainly, a result of a lack of good data due to lack of high quality studies. 5) Some discussion of the following points, if possible, and if not an explanation of why this cannot be done: a) Any evidence of change in quality of submissions of time (within the reports considered) b) Does the poor quality of submissions and subsequent need for clarification delay the NICE decision making process and so partly undermine the rationale for this fast-track STA process? c) Did the clarification letters lead to improvements in the manufacturers' submissions and did this result in a likely change in NICE's decision. In other words, did clarification/improved submission significantly change the results and so the potential decision. This would give a stronger sense of the significance of the quality deficits identified in the paper, but I realise this may be the subject of further research. d) Clarify why the researchers made the recommendations to the manufacturers rather than to NICE, or via NICE to the manufacturers.

REVIEWER	James Raftery Professor of Health Technology Assessment, Wessex Institute, Medical Faculty Southampton University, UK
REVIEW RETURNED	28/11/2011

THE STUDY	As this is a qualitative study, a number of the above questions do not apply. Two "outcomes" are discussed, one to do with identifiying issues and concerns, the second to develop feedback to manufacturers.
GENERAL COMMENTS	This is an interesting paper, one of a small number dealing with the new process that was introduced for NICE's technology appraisal committees in 2005. The authors carry out a qualitative analysis of the first 30 manufacturers submissions and of 21 letters of clarification from the ERGs to the manufacturers. The key finding is that the manufacturers submissions are poor judged against the guidelines specified by NICE. 27/30 were judged to include bias. 17/30 did not perform systematic reviews to the required standard.

	And so on. On the basis of a workshop attended by 50 from the ERGs, plus NICE and NETSCC, 12 recommendations for improvement by manufacturers are formulated which are urged on the manufacturers. My first and major criticism of the paper is that I was left puzzled as whether the findings mattered. Did it make any difference to the Appraisal Committee's recommendation, that the submissions were poor? Since the findings of each of the 30 appraisals are in the public domain I was surprised that the authors did not explore if the better submissions were more successful and vice versa. If all 30 STAs had a positive recommendation for the company involved, then they are hardly likely to heed the recommendations. A company with a poor outcome might want to consider the extent to which this was due to a poor submission. If the technology was potentially cost effective a poor submission leading to negative recommendation would be a serious mistake. But if the technology was highly unlikely to be cost effective by NICE's standards, then a poor biased submission could be deliberate. Ideally, one would want to know the role that the quality of the submission played in relation to both the technology and the recommendation. The authors should at least consider these issues and convince the reader that their findings matter. Second, the paper developed five themes for the ERG report analysis and four categories from the clarification letters. The relationship between these was not clear and was not helped by the account given of how the categories evolved from 8 to 4. Categories and themes seemed to overlap. This should be simplified or explained better. Third, the paper reference s two previous studies (Burls & Sandercock, Miners et al.) but does not consider the extent to which their criticisms of the MTA process apply to the STAs studied in this report. It seemed to me that they probably did, in that they cherry picked comparisons, avoided systematic reviews and were biased. Fourth, this work was funded by the HTA programme which funds the ERGs. The recommendations came from a workshop attended by ERGs, NICE and the HTA programme. Why was industry not invited?. Would the recommendations have been different if they had been present?. The authors should at least consider this. Fifth, it would be helpful to know how many STAs have been done at the time of writing, so one could estimate what proportion this study covered. Finally, some typos such as 'compactor' for 'comparator' (p.5) unless that too was "tongue-incheek".

VERSION 1 – AUTHOR RESPONSE

Comments from Professor Trevor Sheldon

It would have been useful to know if the standard of submissions has been improving over time. Was there for example, any difference between those at the beginning of the process and those submitted further on?

Response: The fact that some of the key themes applied to almost all manufacturer submissions suggests that there was no change over time in this sample. For example, inadequate reporting and the presence of bias in the analyses were both reported for 27 of the 30 submissions. This is perhaps not surprising as the process did not involve any general feedback to manufacturers on their submissions, akin to this study, so any time-related changes would only have been the result of

manufacturers' own prior experience of the process (for those who submitted more than one STA in this sample). There is nothing here to suggest this was the case.

I was somewhat surprised that the researchers had to write these recommendations and send them to the manufacturers; this would appear to be a more appropriate role for NICE. Perhaps, I have misunderstood this.

Response: The remit for this research was to provide feedback and recommendations to manufacturers, the ERGs and NICE on possible ways to improve the STA process. A sentence to clarify this has been added on page 6. This research has been the basis for ongoing discussions between NICE and ABPI on the STA process.

The research would have been strengthened if it has been accompanied by interviews with staff who have carried out the reviews and possibly some in NICE. It could usefully be followed up by research on the perceptions of the manufacturers to better understand how the quality of submissions could be improved.

Response: We agree that the research would have been strengthened if interviews with staff had been undertaken and this was part of the original research proposal. However these were not carried out due to time and resource constraints. We have added a sentence regarding interviewing staff on page 13.

Shorten the objectives para in the Abstract (p3) so as to reduce the text on background – or add a background subheading

Response: We have shortened the objectives paragraph in the abstract.

Add 'government' as a source of pressure for speedier assessment (page 4, line 20)

Response: "Government" has been added.

Change 'compactors' to 'comparators' (page 5, line 42)

Response: "Compactors" has been changed to "comparators".

Possibly better distinguish between bias and uncertainty (page 10, lines 26-36). Uncertainty is also, and often mainly, a result of a lack of good data due to lack of high quality studies.

Response: Bias has been changed to uncertainty.

Some discussion of the following points, if possible, and if not an explanation of why this cannot be done:
a) Any evidence of change in quality of submissions of time (within the reports considered)
Response: This has been dealt with above.

b) Does the poor quality of submissions and subsequent need for clarification delay the NICE decision making process and so partly undermine the rationale for this fast-track STA process?
Response: The poor quality of submissions does not usually delay the process although this does sometimes happen if there is the need for further data requests. The poor quality of submissions does mean that the key issues are not explored fully by the ERGs and therefore by the Appraisal Committee potentially having an impact on the decisions that are made. A sentence to this effect has been added on page 15.

c) Did the clarification letters lead to improvements in the manufacturers' submissions and did this result in a likely change in NICE's decision. In other words, did clarification/improved submission significantly change the results and so the potential decision. This would give a stronger sense of the significance of the quality deficits identified in the paper, but I realise this may be the subject of further research.
Response: c) This is really beyond the scope of this research. We have stated above that the submissions did not appear to get any better through the STAs included in this study. We have added this issue as a point for further research on page 15.

d) Clarify why the researchers made the recommendations to the manufacturers rather than to NICE, or via NICE to the manufacturers.
Response: This point has been dealt with above.

Comments from Professor James Raftery
My first and major criticism of the paper is that I was left puzzled as whether the findings mattered. Did it make any difference to the Appraisal Committee's recommendation, that the submissions were poor?

Response: This has been covered above.

Since the findings of each of the 30 appraisals are in the public domain I was surprised that the authors did not explore if the better submissions were more successful and vice versa. If all 30 STAs had a positive recommendation for the company involved, then they are hardly likely to heed the recommendations. A company with a poor outcome might want to consider the extent to which this was due to a poor submission. If the technology was potentially cost effective a poor submission leading to negative recommendation would be a serious mistake. But if the technology was highly unlikely to be cost effective by NICE's standards, then a poor biased submission could be deliberate. Ideally, one would want to know the role that the quality of the submission played in relation to both the technology and the recommendation. The authors should at least consider these issues and convince the reader that their findings matter.

Response: This was beyond the remit of this research. We have covered the point about whether or not poor submissions matter above.

Second, the paper developed five themes for the ERG report analysis and four categories from the clarification letters. The relationship between these was not clear and was not helped by the account given of how the categories evolved from 8 to 4. Categories and themes seemed to overlap. This should be simplified or explained better.

Response: This has now been changed and the clarification letter analysis precedes the ERG report analysis with an explanation of the relationship between the two on page 11.

Third, the paper reference s two previous studies (Burls & Sandercock, Miners et al.) but does not consider the extent to which their criticisms of the MTA process apply to the STAs studied in this report. It seemed to me that they probably did, in that they cherry picked comparisons, avoided systematic reviews and were biased

Response: We feel this point has already been covered on page 5.

Fourth, this work was funded by the HTA programme which funds the ERGs. The recommendations came from a workshop attended by ERGs, NICE and the HTA programme. Why was industry not

invited? Would the recommendations have been different if they had been present? The authors should at least consider this.

Response: We feel this has been covered in previous points.

Fifth, it would be helpful to know how many STAs have been done at the time of writing, so one could estimate what proportion this study covered.

Response: At the time of writing, the ERG reports for 94 STAs have been completed, although not all of these will have had NICE guidance issued yet. We have added a statement on page 14.

Finally, some typos such as ‘compactor’ for ‘comparator’ (p.5) unless that too was “tongue-in-cheek”.

Response: This has been changed.

We hope these changes adequately address the referees' comments. Please do not hesitate to contact us if you require any further information or changes.

Thank you for considering this paper for publication in BMJ Open.

Yours sincerely,
Eva Kaltenthaler

VERSION 2 – REVIEW

REVIEWER	Professor Trevor A Sheldon Deputy Vice-Chancellor University of York I am also chair of the Board of the York Health Economics Consortium (a wholly owned subsidiary of the University) which conducts research work for both NICE and industry
REVIEW RETURNED	14/12/2011

RESULTS & CONCLUSIONS	See my comments below about the implications for the STA process and NICE
GENERAL COMMENTS	I am content with the revisions made and think that the paper is ready for publication and will provide a useful window of the Single Technology Assessment Process. Just two related points which might be addressed relatively easily. The authors sate that: “It is hoped that uptake of these recommendations suggestions by manufacturers will result in more transparent and internally consistent submissions and will improve the efficiency of the current NICE STA process and any subsequent process.” However, given that it appears from the response to the referees’ comments that most of these technologies get approved by NICE despite poor submissions and that there is little delay in this process when asking for clarifications, I cannot see where the incentive is to implement the suggestions. So in addition to the text: “The poor quality of submissions does not usually delay the process although this does sometimes occur if there is the need for further data requests. The poor quality of

	submissions does mean that the key issues may not be explored fully by the ERGs and therefore by the Appraisal Committee, potentially having an impact on the decisions that are made." I would suggest they point out to readers that this demonstrates (a) that the current STA process as described is flawed and likely to result in some poor decisions for the NHS and society and (b) that NICE should require a higher quality of submission before approving the funding of technologies.

REVIEWER	James Raftery Professor of HTA University of Southampton, UK No conflicts of interest
REVIEW RETURNED	21/12/2011

THE STUDY	The authors response to most of the two referees comments is along the lines of "not in remit". However the remit of the study is not clearly stated. The nearest to such a statement is "The purpose of this research study was to identify common issues and concerns identified by the ERGs in their analyses of MS and to use these as a basis to provide feedback to manufacturers to assist them in future submission development as well as recommendations to the ERGs and NICE." This should be strengthened to clarify what is within and without remit.
GENERAL COMMENTS	I don't think the authors have addressed the comments made with any enthusiasm. The revisions are minor but the article is worth publishing.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1 comment:
So in addition to the text: "The poor quality of submissions does not usually delay the process although this does sometimes occur if there is the need for further data requests. The poor quality of submissions does mean that the key issues may not be explored fully by the ERGs and therefore by the Appraisal Committee, potentially having an impact on the decisions that are made." I would suggest they point out to readers that this demonstrates (a) that the current STA process as described is flawed and likely to result in some poor decisions for the NHS and society and (b) that NICE should require a higher quality of submission before approving the funding of technologies.

Response:
We do not feel that the current STA process is flawed so have not added this statement. However we have revised the statement regarding the delay in the process on page 15.

Reviewer 2 comment:
The authors response to most of the two referees comments is along the lines of "not in remit". However the remit of the study is not clearly stated. The nearest to such a statement is "The purpose of this research study was to identify common issues and concerns identified by the ERGs in their analyses of MS and to use these as a basis to provide feedback to manufacturers to assist them in future submission development as well as recommendations to the ERGs and NICE." This should be strengthened to clarify what is within and without remit.

Response:
A sentence regarding the remit of this research has been added to page 14. Those areas not dealt with in this paper are suggested future research areas.

We hope these address the comments made by the reviewers. Thank you once again for considerding this paper for BMJ Open.

Yours sincerely,
Eva Kaltenthaler